The University of Chicago

CONDITIONS FOR A MINIMUM OF A FUNCTIONAL

A DISSERTATION SUBMITTED TO THE FACULTY OF THE DIVISION OF THE PHYSICAL SCIENCES IN CANDIDACY FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF MATHEMATICS
1936

By
HERMAN HEINE GOLDSTINE

Private Edition, Distributed by
THE UNIVERSITY OF CHICAGO LIBRARIES
CHICAGO, ILLINOIS

Reprinted from
CONTRIBUTIONS TO THE CALCULUS OF VARIATIONS, 1933–37
The University of Chicago Press, 1937

The University of Chicago

CONDITIONS FOR A MINIMUM OF A FUNCTIONAL

A DISSERTATION SUBMITTED TO THE FACULTY OF THE DIVISION OF THE PHYSICAL SCIENCES IN CANDIDACY FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF MATHEMATICS
1936

By
HERMAN HEINE GOLDSTINE

Private Edition, Distributed by
THE UNIVERSITY OF CHICAGO LIBRARIES
CHICAGO, ILLINOIS

Reprinted from
CONTRIBUTIONS TO THE CALCULUS OF VARIATIONS, 1933–37
The University of Chicago Press, 1937

R. UNIV.
BIBLIOTHEEK
LEIDEN

TABLE OF CONTENTS

Section Page

Introduction 1

CHAPTER I

PROPERTIES OF CONTINUOUS BILINEAR FUNCTIONALS

Introduction 3

1. Definitions 3

2. Properties of continuous bilinear functionals 5

3. An integral equation 12

CHAPTER II

FIRST AND SECOND DIFFERENTIALS

Introduction 16

4. Definitions and first properties 16

CHAPTER III

THE MINIMA OF FUNCTIONALS

Introduction 20

5. The statement of the problem 20

6. A representation of J and K 21

7. The necessary conditions 22

8. The sufficient conditions 25

9. Excursus 26

CHAPTER IV

APPLICATIONS OF THE GENERAL THEORY

Introduction 32

10. The calculus of variations 32

11. A non-calculus of variations example 37

Digitized by the Internet Archive
in 2026

https://archive.org/details/IA41548411_0099

CONDITIONS FOR A MINIMUM OF A FUNCTIONAL

INTRODUCTION

A study of the extremizing properties of numerically valued functions is of fundamental importance to a theory of such functions and has been treated at great length in the calculus of variations. However, the functionals which are discussed in this branch of analysis are somewhat special in character; hence their treatment by no means exhausts the theory of the minima of functionals.

With the inception of the general theory of functionals considerable impetus was given to this problem with the result that a number of attempts was made to solve it.[1] The most complete of these was made by Le Stourgeon who obtained two conditions which are necessary for a minimum; however, the generality of her results is marred by the explicit assumption of certain stringent hypotheses and the tacit use of others.[2]

In this paper we propose to find necessary conditions that a functional, having a first and second differential in a given region, be a minimum at a point; moreover we show that by strength-

---

[1] An historical survey of the problem and an extensive bibliography of the literature which arose in this connection were made by R. G. Sanger, "Functions of lines and the calculus of variations," Contributions to the Calculus of Variations 1931-1932, Chicago, 1933, pp. 191-293.

[2] E. Le Stourgeon, Minima of functions of lines, Transactions of the American Mathematical Society, vol. 21 (1920), pp. 357-383. The aforementioned assumptions are made on pp. 369 f. and 376 ff.

ening these conditions, as is done in the calculus of variations, we obtain a sufficiency theorem. No restrictions are imposed upon the functional except that it have a second differential with certain continuity properties.

The first chapter comprises a study of those properties of continuous bilinear functionals which are fundamental to this solution of the problem. Chapter II is a statement of the definitions of differentials and a few of their immediate consequences. In the third chapter we find necessary conditions and sufficient conditions that a functional having certain differentiability properties have a minimum in a given class. Chapter IV is devoted to a detailed examination of two special cases of the general theory developed in the previous chapter. One of these cases is the fixed end-point problem of the calculus of variations; the other is a problem not of the calculus of variations type.

CHAPTER I

PROPERTIES OF CONTINUOUS BILINEAR FUNCTIONALS

Introduction. In this chapter we shall be concerned mainly with functionals defined on the class of continuous functions or on subspaces of this class. In the first sections can be found a representation theorem for continuous bilinear functionals on certain subsets of the space of continuous functions. The chapter closes with a section devoted to a study of an integral equation and its relation to the positive properties of quadratic functionals. The major purpose of the chapter is to relate the characteristic values of an integral equation to certain bilinear functionals; this relation is indispensable to the study of the second differential.

1. Definitions. For the purposes of this and succeeding chapters we shall define three spaces, $\mathfrak{C}$, $\mathfrak{D}$, and $\mathfrak{N}$: the first will be understood to be the class of all real-valued functions ξ defined and continuous on the interval from 0 to 1; the second will be the class of all δ in $\mathfrak{C}$ which are subject to the restriction that

$$\int_0^1 \delta(t)dt = 0;$$

lastly, $\mathfrak{N}$ is the set of all ν such that ν is in $\mathfrak{D}$ and has a continuous derivative on the interval from 0 to 1. These spaces will have associated with them as a norm[1] the maximum of the abso-

---

[1] See S. Banach, Théorie des Opérations Linéaires, Warsaw, 1932, p. 53. We symbolize the norm of a function ξ by $\|\xi\|$.

lute value of the function under consideration.

We mean by J on $\mathfrak{C}$ to $\mathfrak{A}$ or K on $\mathfrak{C}\mathfrak{C}$ to $\mathfrak{A}$ a function which makes correspond to every ξ or to every pair (ξ_1, ξ_2), respectively, a real number a.[1] A functional J on $\mathfrak{C}$ to $\mathfrak{A}$ is said to be linear in the event that $J(\xi_1) + J(\xi_2) = J(\xi_1 + \xi_2)$ for every ξ_1 and ξ_2. A bilinear functional K on $\mathfrak{C}\mathfrak{C}$ to $\mathfrak{A}$ is one which is linear in each of its arguments. Such functionals are continuous in case the fact that

$$\lim_{n=\infty} \|\xi_n - \xi\| = 0 \quad \text{or} \quad \lim_{n=\infty} \|\xi_{ni} - \xi_i\| = 0 \quad (i = 1, 2)$$

implies that

$$\lim_{n=\infty} J(\xi_n) = J(\xi) \quad \text{or} \quad \lim_{m,n=\infty} K(\xi_{m1}, \xi_{n2}) = K(\xi_1, \xi_2).$$

It can easily be shown that linear and bilinear functions which are continuous are also modular, i.e., there are two numbers a_1 and a_2 such that

$$(1{:}1) \quad |J(\xi_1)| \leqq a_1\|\xi_1\| \text{ and } |K(\xi_1, \xi_2)| \leqq a_2\|\xi_1\|\cdot\|\xi_2\|,$$

for every ξ_1 and ξ_2; the converse is also true.

Fréchet has proved[2] that, for every continuous bilinear function K on $\mathfrak{C}\mathfrak{C}$ to $\mathfrak{A}$, there exists a function k(x, y) of limited variation in the sense of Fréchet[3] with the following properties:

$$(1{:}2) \qquad K(\xi_1, \xi_2) = \int_0^1 \int_0^1 \xi_1(x)\, \xi_2(y) d_x d_y k(x, y),$$

---

[1] We shall make use of this notation for spaces other than $\mathfrak{C}$; the space $\mathfrak{A}$ is to be understood to be the set of all real numbers.

[2] M. Fréchet, Sur les fonctionelles bilinéaires, Transactions of the American Mathematical Society, vol. 16 (1915), pp. 215-244.

[3] Ibid., pp. 223 ff.

for every ξ_1 and ξ_2; the smallest effective a_2 in the second inequality of (1:1) is the Fréchet total variation[1] of k; and $k(1, y) = k(x, 1) = 0$ for $0 \leqq x \leqq 1$ and $0 \leqq y \leqq 1$.

2. Properties of continuous bilinear functionals. We turn our attention to the consideration of a number of lemmae which will be basic in the succeeding pages. The first of these to be considered is

LEMMA 2:1. *If K on $\mathfrak{C}\mathfrak{C}$ to $\mathfrak{A}$ is bilinear and continuous, and if*

$$K(\xi_1, \xi_2) = K(\xi_2, \xi_1),$$

for every ξ_1 and ξ_2, then there exists a symmetric function k of limited variation in the sense of Fréchet such that

$$K(\xi_1, \xi_2) = \int_0^1 \int_0^1 \xi_1(x)\, \xi_2(y) d_x d_y k(x, y),$$

for every choice of ξ_1 and ξ_2; moreover

$$k(x, 1) = k(1, y) = 0 \qquad (0 \leqq x \leqq 1,\ 0 \leqq y \leqq 1).$$

It follows readily from the definition of the double Stieltjes integral[2] and equation (1:2) that

$$K(\xi_1, \xi_2) = \int_0^1 \int_0^1 \xi_1(x)\, \xi_2(y) d_x d_y k_0(x, y)$$

$$= \int_0^1 \int_0^1 \xi_1(x)\, \xi_2(y) d_x d_y [k_0(x, y) + k_0(y, x)]/2;$$

hence, if we define $k(x, y)$ to be $[k_0(x, y) + k_0(y, x)]/2$, it is effective in the lemma. Lastly, recalling that $k_0(x, 1) = k_0(1, y) = 0$, we note that $k(x, 1) = k(1, y) = 0$.

---

[1] Ibid., pp. 219 f.

[2] Ibid., pp. 235 ff.

We now define a function that vanishes at every point on the boundary of the square.[1]

$$k(x, y) \equiv k_1(x, y) + k_1(0, 0)(1 - x)(1 - y) - k_1(x, 0)(1 - y) - k_1(0, y)(1 - x), \tag{2:1}$$

where k_1 is the function whose existence is asserted in the last lemma. We summarize the properties of k in the following lemma:

LEMMA 2:2. The function k defined by equation (2:1) is symmetric, bounded, of limited variation in the sense of Fréchet, and zero on the boundary of the square.

The symmetry of k and its vanishing on the boundary of the square are immediate consequences of equation (2:1) and Lemma 2:1. Next we note that k is of limited variation in the sense of Fréchet[2] since each term in the right member of equation (2:1) is of limited variation; the first two terms obviously have this property while the last two have this property since $k_1(x, 1)$ and $k_1(1, y)$ vanish.[3]

Lastly we see that k is bounded because

$$|k(x, y)| = |k(1, 1) - k(1, y) - k(x, 1) + k(x, y)| \leqq (\text{Fréchet total variation of } k),$$

for every pair (x, y) in the square.

An immediate consequence of this lemma is

LEMMA 2:3. For every continuous, bilinear, and symmetric functional K on $\mathfrak{C}$ to $\mathfrak{A}$, there exists a function k, which has

---

[1] By the square we mean the class of all pairs (x, y) such that $0 \leqq x \leqq 1$, $0 \leqq y \leqq 1$.

[2] The sum of two functions of limited variation in the sense of Fréchet evidently also has this property.

[3] Fréchet, op. cit., pp. 229 f.

the properties stated in Lemma 2:2, such that

$$K(\delta_1, \delta_2) = \int_0^1 \int_0^1 \delta_1(x)\, \delta_2(y) d_x d_y k(x, y)$$

for every δ_1 and δ_2 in $\mathfrak{D}$.

If k_1 be the kernel[1] of K and if k be defined by equation (2:1), then it can easily be shown that

$$\int_0^1 \int_0^1 \delta_1(x)\, \delta_2(y) d_x d_y k(x, y) \tag{2:3}$$

is continuous, bilinear, and symmetric.[2]

To show that this function k is effective in the lemma we note that expression (2:3) is equal to

$$\int_0^1 \int_0^1 \delta_1(x)\, \delta_2(y) d_x d_y k_1(x, y)$$

$$+ \int_0^1 \int_0^1 \delta_1(x)\, \delta_2(y) d_x d_y k_1(0, 0)(1 - x)(1 - y)$$

$$- \int_0^1 \int_0^1 \delta_1(x)\, \delta_2(y) d_x d_y k_1(x, 0)(1 - y)$$

$$- \int_0^1 \int_0^1 \delta_1(x)\, \delta_2(y) d_x d_y k_1(0, y)(1 - x);$$

that the second term vanishes because it is equal to

$$k_1(0, 0) \cdot \int_0^1 \delta_1(x) dx \cdot \int_0^1 \delta_2(y) dy;$$

and that the last two terms vanish for δ_1 and δ_1 in $\mathfrak{D}$ since they are equal to

---

[1] The kernel of a continuous, bilinear, and symmetric function on $\mathfrak{CC}$ to $\mathfrak{A}$ is the function whose existence is asserted in Lemma 2:1.

[2] Fréchet, op. cit., p. 227.

$$-\int_0^1 \delta_2(y)dy \cdot \int_0^1 \delta_1(x)d_x k_1(x, 0),$$

$$-\int_0^1 \delta_1(x)dx \cdot \int_0^1 \delta_2(y)d_y k_1(0, y),$$

respectively. Whence, it follows that $K(\delta_1, \delta_2)$ is expressible by means of expression (2:3).

In the future we shall be concerned with functionals of type $\mathfrak{R}$: they are, by definition, continuous, bilinear, and symmetric functionals on $\mathfrak{CC}$ to $\mathfrak{A}$ which have kernels continuous in both variables together. In terms of this definition we state

THEOREM 1. *If K is a continuous, bilinear, and symmetric functional on $\mathfrak{CC}$ to $\mathfrak{A}$, then there exists a symmetric, bounded, and Riemann integrable function k(x, y), vanishing on the boundary of the square, such that*

$$K(\nu_1, \nu_2) = \int_0^1 \int_0^1 \nu_1'(x) k(x, y) \nu_2'(y)\, dx\, dy,$$

for every ν_1, ν_2 in $\mathfrak{N}$, where ν' is the derivative of ν. Moreover if K is of type $\mathfrak{R}$, then there is a unique continuous function satisfying the equation above.

By the last lemma there is a bounded function k which is symmetric, vanishes on the boundary of the square, and is of limited variation in the sense of Fréchet such that

$$K(\nu_1, \nu_2) = \int_0^1 \int_0^1 \nu_1(x)\, \nu_2(y) d_x d_y k(x, y). \tag{2:4}$$

It will now be shown that this function k is effective in the theorem.

Consider any two partitions of the interval from 0 to 1 whose partitioning points x_i and t_j, respectively, ($i = 0, 1, \ldots, m$; $j = 0, 1, \ldots, n$) are such that $0 = x_0 = t_0$ and $1 = x_m = t_n$;

consider also two sets of numbers ρ_i and σ_j ($i = 1, 2, \ldots, m$; $j = 1, 2, \ldots, n$) such that $x_{i-1} \leqq \rho_i \leqq x_i$ and $t_{j-1} \leqq \sigma_j \leqq t_j$ for $i = 2, \ldots, m-1$ and $j = 2, \ldots, n-1$, and $0 = \rho_1 = \sigma_1$ and $1 = \rho_m = \sigma_n$. Then since k vanishes on the boundary of the square, we note the algebraic identity

$$\sum_{i=1}^{m}\sum_{j=1}^{n} \rho_i \sigma_j [k(x_i, t_j) - k(x_i, t_{j-1}) - k(x_{i-1}, t_j) + k(x_{i-1}, t_{j-1})]$$

$$= \sum_{i=1}^{m-1}\sum_{j=1}^{n-1} k(x_i, t_j)(\rho_{i+1} - \rho_i)(\sigma_{j+1} - \sigma_j), \tag{2:5}$$

for every integer m and n.

However, since the limit, as the norms of the partitions simultaneously approach zero, of the first member of equation (2:5) exists and is equal to

$$\int_0^1 \int_0^1 xy\, d_x d_y k(x, y),$$

the limit, as the largest of the $\rho_{i+1} - \rho_i$ and $\sigma_{j+1} - \sigma_j$ approach zero, of the second member exists; but when this limit exists independently of the choice of the partitions and of the points (x_i, t_j) it is, by definition, the double Riemann integral of k.[1] Hence we have shown that k is Riemann integrable. In an analogous fashion we can show the existence of

$$\lim_{ND=0} \sum_{i=1}^{m}\sum_{j=1}^{n} k(\rho_i, \sigma_j)[\nu_1(x_i) - \nu_1(x_{i-1})][\nu_2(y_j) - \nu_2(y_{j-1})]$$

$$= \int_0^1 \int_0^1 \nu_1(x)\, \nu_2(y)\, d_x d_y k(x, y), \tag{2:6}$$

---

[1] More generally, we can prove by analogous methods that any function of limited variation in the sense of Fréchet which is of limited variation in y for a particular x and of limited variation in x for a particular y is Riemann integrable.

where by ND we mean the norms of the partitions formed by the points ρ_i and σ_j; moreover it is evident that $\nu_1'(x)k(x, y)\nu_2'(y)$ is Riemann integrable. Hence, if we take any two partitions of the interval from 0 to 1 of the kind described above and choose two sets of numbers ρ_i and σ_j such that, for $i = 1, \ldots, m$ and $j = 1, \ldots, n$, it is true that[1]

$$\nu_1(x_i) - \nu_1(x_{i-1}) = \nu_1'(\rho_i)(x_i - x_{i-1}),$$
$$\nu_2(y_j) - \nu_2(y_{j-1}) = \nu_2'(\sigma_j)(y_j - y_{j-1}),$$

$x_{i-1} < \rho_i < x_i$, and $y_{j-1} < \sigma_j < y_j$, then

$$\int_0^1\int_0^1 \nu_1'(x)k(x, y)\,\nu_2'(y)\,dx\,dy$$
$$= \lim_{ND=0} \sum_{i=1}^{m}\sum_{j=1}^{n} \nu_1'(\rho_i)k(\rho_i, \sigma_j)\,\nu_2'(\sigma_j)(x_i - x_{i-1})(y_j - y_{j-1})$$
$$= \lim_{ND=0} \sum_{i=1}^{m}\sum_{j=1}^{n} k(\rho_i, \sigma_j)[\nu_1(x_i) - \nu_1(x_{i-1})][\nu_2(y_j) - \nu_2(y_{j-1})]$$
$$= \int_0^1\int_0^1 \nu_1(x)\,\nu_2(y)d_xd_yk(x, y)$$

by virtue of equation (2:6).

To establish the latter half of the theorem we first observe that equation (2:4) and the fact that K is of type $\mathfrak{N}$ imply the existence of a continuous function k satisfying the first part of the theorem; hence it remains only to show that k is unique.

To establish this fact we suppose that there are two continuous functions k_1 and k_2 which are effective in the theorem; therefore if k_0 be defined to be $k_1 - k_2$, we have at once that k_0

[1] The existence of these two sets is a consequence of the definition of $\mathfrak{N}$ and of the law of the mean for derivatives.

is continuous and that

$$\int_0^1\int_0^1 \nu_1'(x)k_0(x, y)\,\nu_2'(y)\,dx\,dy = 0, \tag{2:7}$$

for every ν_1 and ν_2 in $\mathfrak{N}$. For $0 \leqq s \leqq 1$ let ν_1 and ν_2 be defined by the relations

$$\nu_1(x) \equiv \int_0^x k_0(x, t)dt - \int_0^1\int_0^x k_0(x, t)\,dt\,dx,$$

$$h(s, y) \equiv \int_0^1 k_0(s, t)k_0(t, y)dt,$$

$$\nu_2(y) \equiv \int_0^y h(s, u)du - \int_0^1\int_0^y h(s, u)\,du\,dy.$$

Whence equation (2:7) becomes

$$\int_0^1 \nu_2'(y)\int_0^1 \nu_1'(x)k_0(x, y)\,dx\,dy$$

$$= \int_0^1 h^2(s, y)dy = 0 \qquad (0 \leqq s \leqq 1);$$

thus we see that $h(s, y) = 0$ for $0 \leqq s \leqq 1$ and $0 \leqq y \leqq 1$. In particular we note that

$$h(s, s) = \int_0^1 k_0^2(s, t)dt = 0 \qquad (0 \leqq s \leqq 1),$$

from which it follows that $k_0(s, t)$ is identically zero.

As a conclusion to this section we state a result on the positive properties of continuous bilinear functionals.

THEOREM 2. *A necessary and sufficient condition that a continuous bilinear function K on $\mathfrak{C}\mathfrak{C}$ to $\mathfrak{A}$ be such that $K(\delta, \delta) \geqq 0$, for every δ in $\mathfrak{D}$, is that $K(\nu, \nu) \geqq 0$, for every ν in $\mathfrak{N}$.*

If $K(\delta, \delta) \geqq 0$ for every δ, it is obvious that $K(\nu, \nu)$

$\geqq 0$ since $\mathfrak{N}$ is in $\mathfrak{D}$. Hence it remains only to show the truth of the converse.

Suppose that $K(\nu, \nu) \geqq 0$ for ν in $\mathfrak{N}$ and that there is a δ_0 such that $K(\delta_0, \delta_0) < 0$. Then by the Weierstrass approximation theorem[1] there is a sequence $\{\xi_n\}$ of polynomials which converge uniformly to δ_0; as a consequence we note that

$$\lim_{n=\infty} \int_0^1 \xi_n(t)dt = 0.$$

Hence it is apparent that the sequence $\{\nu_n\}$ defined by the identity

$$\nu_n \equiv \xi_n - \int_0^1 \xi_n(t)dt \qquad (n = 1, 2, \ldots)$$

also converges uniformly to δ_0, and that every ν_n is in $\mathfrak{N}$.

Whence, as a result of the continuity of K, it must be the case that $K(\nu_n, \nu_n) < 0$, for sufficiently large integers n; but this is a contradiction of our hypothesis.

3. An integral equation. In this section a necessary and sufficient condition that a functional K of type $\mathfrak{R}$ be such that $K(\nu, \nu) \geqq 0$ for any ν is sought. To find such a condition we consider the integral equation

$$\int_0^1 k(x, y)\xi(y)dy = \lambda \xi(x), \tag{3:1}$$

where k is a symmetric function which is continuous in x and y together and not identically zero in the square; we seek solutions ξ in $\mathfrak{C}$ which are not identically zero. Then we know that there

---

[1] See E. W. Hobson, The Theory of Functions of a Real Variable, Cambridge, 1926, vol. 2, pp. 228 ff.

exists[1] a characteristic value $\lambda \neq 0$; that the class of all characteristic values is, at most, denumerable; that every characteristic value is real; and that if $\{\xi_i\}$ be a normed and orthogonal basis[2] for the integral equation (3:1), then

$$\int_0^1 \int_0^1 \gamma'(x)k(x, y)\gamma'(y)\,dx\,dy = \sum_i \lambda_i \left[\int_0^1 \gamma'(t)\cdot \xi_i(t)dt\right]^2, \tag{3:2}$$

where the sum is taken over all ξ_i in the basis. Moreover we see that

LEMMA 3:1. *For every characteristic solution ξ of the integral equation (3:1) there exists a γ in $\mathfrak{N}$ such that $\gamma' = \xi$.*

If γ be defined by the identity

$$\gamma(x) \equiv \int_0^x \xi(t)dt - \int_0^1 \int_0^x \xi(t)\,dt\,dx,$$

then it is obviously effective in the lemma. Hence the problem of finding numbers λ and non-identically zero functions γ in $\mathfrak{N}$ such that

$$\int_0^1 k(x, y)\,\gamma'(y)dy = \lambda\,\gamma'(x)$$

is equivalent to solving equation (3:1). This leads us to state the fundamental theorem of this chapter. To simplify its statement we introduce the term reduced kernel over $\mathfrak{N}$ which is, by

---

[1] See E. Goursat, Cours d'Analyse Mathématique, Paris, 1927, vol. 3, pp. 439 ff.

[2] The sequence is a basis in the sense that every ξ_i is a solution of (3:1), and any solution is expressible as a linear combination of a finite number of elements of $\{\xi_i\}$; moreover if λ_i be the characteristic value associated with ξ_i, then $|\lambda_i| \geqq |\lambda_{i+1}|$.

definition, the function whose existence is asserted in Theorem 1.

THEOREM 3. A necessary and sufficient condition that a functional K of type $\mathfrak{R}$, which is not zero on $\mathfrak{D}$, be such that $K(\delta, \delta) \geqq 0$, for every δ in $\mathfrak{D}$, is that there is no negative characteristic value for the integral equation

$$\int_0^1 k(x, y)\,\xi(y)dy = \lambda\,\xi(x), \tag{3:3}$$

where k is the reduced kernel over $\mathfrak{N}$ of K.

If $K(\delta, \delta) \geqq 0$ for every δ and if λ_1 is any characteristic value[1] of equation (3:3), then

$$\begin{aligned} K(\nu_1, \nu_1) &= \int_0^1\int_0^1 \nu_1'(x)k(x, y)\,\nu_1'(y)\,dx\,dy \\ &= \lambda_1\int_0^1 [\nu_1'(t)]^2 dt \geqq 0, \end{aligned} \tag{3:4}$$

where ν_1 has been defined in terms of ξ_1, a characteristic solution of equation (3:3) associated with λ_1, as follows

$$\nu_1(x) \equiv \int_0^x \xi_1(t)dt - \int_0^1\int_0^x \xi_1(t)\,dt\,dx;$$

but since ν' is not identically zero, $\lambda_1 \geqq 0$.

Conversely suppose that there is no negative characteristic value. Then it follows from equation (3:2) that $K(\nu, \nu) \geqq 0$ since no term in the right member of equation (3:2) is negative. Consequently by Theorem 2, $K(\delta, \delta) \geqq 0$.

If we suppose that every characteristic value is positive, then the only function orthogonal to every function in the basis

---

[1] It follows from the hypothesis that k is not identically zero and hence has a non-zero characteristic value.

for equation (3:3) is the zero function.[1] Hence it is evident from equation (3:2) that $K(\nu, \nu) \neq 0$ whenever ν is not the zero function; consequently we are led to state the following

COROLLARY 1. $K(\nu, \nu)$ *is positive, for* ν *in* $\mathfrak{N}$ *and not the zero function, if and only if every characteristic value of the integral equation* (3:3) *is positive.*

Suppose that there is a negative or zero characteristic value, λ_1. Then we see at once that

$$K(\nu_1, \nu_1) = \lambda_1 \int_0^1 [\nu_1'(t)]^2 dt \leqq 0,$$

where ν_1 has been defined in terms of ξ_1, a characteristic solution corresponding to λ_1, by means of the equation immediately succeeding equation (3:4); however this is clearly impossible.

---

[1] See Goursat, *op. cit.*, pp. 445 f.

CHAPTER II

FIRST AND SECOND DIFFERENTIALS

Introduction. In the main this chapter contains the definitions of first and second differentials and a few of their properties. It will be shown that the second differential is bilinear and continuous. Hence the results of the previous chapter are applicable to the differentials of functions defined on $\mathfrak{E}$ or linear subspaces of it.

4. Definitions and first properties. In this section we shall indicate by $\mathfrak{X}$ a Banach space, i.e., a normed linear space which is complete;[1] further by the d-neighborhood of a point x_0 (in symbols, $N_d(x_0)$) we shall mean the set of all x such that $\|x - x_0\| < d$. In terms of these definitions one says that a region of $\mathfrak{X}$ is a set $\mathfrak{X}_0$ of points of $\mathfrak{X}$ such that every x_0 in $\mathfrak{X}_0$ has a d-neighborhood which contains only points of $\mathfrak{X}_0$.

A functional I on $\mathfrak{X}_0$ to $\mathfrak{Y}$ has a first differential at a point[2] x_0 in case there exists a continuous linear functional J on $\mathfrak{X}$ to $\mathfrak{Y}$ such that, for every $e > 0$, there is an $N_d(x_0)$ with the property that

$$|I(x) - I(x_0) - J(x - x_0)| \leqq e\,\|x - x_0\|,$$

whenever x is in $\mathfrak{X}_0$ and $N_d(x_0)$. Moreover I has a second differ-

---

[1] See L. M. Graves, Topics in the functional calculus, Bulletin of the American Mathematical Society, vol. 41 (1935), p. 645.

[2] T. H. Hildebrandt and L. M. Graves, Implicit functions and their differentials in general analysis, Transactions of the American Mathematical Society, vol. 29 (1927), p. 165.

ential at x_0 when the following conditions are satisfied:

(1) there exists an $N_d(x_0)$ such that I has a first differential J on $N_d(x_0)\mathfrak{X}$ to $\mathfrak{Y}$, i.e., I has a differential at every point of $N_d(x_0)$.

(2) for every x_1 in $\mathfrak{X}$, it is true that J has a first differential $K(x_0; x_1, x_2)$ at x_0.

A systematic study of these concepts can be found in a joint paper of Hildebrandt and Graves[1] and in an article by Graves.[2] We shall content ourselves with proving a few miscellaneous properties which will be useful for our purposes.

LEMMA 4:1. *If I on $\mathfrak{X}_0$ to $\mathfrak{Y}$ has a first or second differential at a point x_0, this differential is unique and expressible as*

$$\frac{d}{da} I(x_0 + ax)\Big|^{a=0} \quad \text{or} \quad \frac{d^2}{da^2} I(x_0 + ax)\Big|^{a=0},$$

respectively.

For a proof the reader is referred to the paper of Hildebrandt and Graves cited above.[3]

LEMMA 4:2. *The second differential K on $\mathfrak{X}\mathfrak{X}$ to $\mathfrak{Y}$ is bilinear, continuous, and symmetric.*[4]

Since Graves showed in detail that K is symmetric,[5] it suffices to establish only the remainder of the lemma. To do this, it will first be shown that K is continuous in its first argument

---

[1] Ibid., pp. 135 ff.

[2] Graves, loc. cit., pp. 648-652.

[3] Hildebrandt and Graves, loc. cit., p. 136.

[4] Ibid., p. 651.

[5] L. M. Graves, Riemann Integration and Taylor's theorem in general analysis, Transactions of the American Mathematical Society, vol. 29 (1927), pp. 176 f.

x_1. It should be noted that the bilinearity of K is immediate.

By the previous lemma and the definition of K it is clear that

$$K(x_1, x_2) = \lim_{a=0} [J(x_0 + ax_2, x_1) - J(x_0, x_1)]/a,$$

for every x_1 and x_2. Hence defining $g_n(x_1)$ to be

$$n[J(x_0 + x_2/n, x_1) - J(x_0, x_1)],$$

we observe that, for every n and x_2, g_n is a linear and continuous functional on $\mathfrak{X}$ to $\mathfrak{Y}$; moreover from the definition of g_n it is immediate that

$$\lim_{n=\infty} g_n(x_1) = K(x_1, x_2),$$

for each x_2. Hence the sequence $\{f_n\} \equiv \{|g_n|\}$ satisfies the hypotheses of Hildebrandt's theorem[1] since $\mathfrak{X}$ is complete; as a consequence there exists a number $a(x_2)$ such that $f_n(x_1) \leqq a(x_2)\|x_1\|$, for every n, x_1, and x_2. Therefore $|K(x_1, x_2)| \leqq a(x_2)\|x_1\|$, which establishes the continuity of K in its first argument.

Since K is continuous in x_1 and also in x_2, we again make use of Hildebrandt's theorem to show that it is continuous in both variables together. Suppose that K is not continuous (or what is the same thing, modular) in both variables together. Then for every integer n, there exists an x_{1n} and x_{2n}, whose norms are unity, such that $|K(x_{1n}, x_{2n})| > n$. But if $g_n(x_2)$ be defined as $|K(x_{1n}, x_2)|$, it satisfies the hypotheses of Hildebrandt's

---

[1] T. H. Hildebrandt, On the uniform limitedness of sets of functional operations, Bulletin of the American Mathematical Society, vol. 29 (1923), pp. 309 ff.

theorem; consequently there exists a number a for which $g_n(x_s) \leqq a \|x_s\|$, for every n and x_s. Therefore we have $n < g_n(x_{sn}) \leqq a$, for every integer n, which is impossible.

CHAPTER III

THE MINIMA OF FUNCTIONALS

Introduction. The theorems of the previous sections will be utilized in this chapter to establish three conditions which are necessary for a certain functional to have a minimum and two conditions which are sufficient. The first two sections are auxiliary to this purpose: in particular the first section is a formal statement of the problem, and the second deals with the representation of the differentials by means of definite integrals. The necessary conditions are deduced in Section 8 and the sufficient ones in the succeeding section.

5. The statement of the problem. The basis for this theory will be the spaces $\mathfrak{C}$, $\mathfrak{D}$, and $\mathfrak{N}$, a region $\mathfrak{C}_0$ of $\mathfrak{C}$, a fixed element ξ_0 of $\mathfrak{C}_0$; a subset $\mathfrak{B}$ of $\mathfrak{C}_0$, which is defined to be the set of all elements ζ of $\mathfrak{C}_0$ such that $\zeta - \xi_0$ lies in $\mathfrak{D}$; and a functional I on $\mathfrak{C}_0$ to $\mathfrak{A}$ with the following property:

(H. 1) I has a second differential $K(\zeta_0; \xi_1, \xi_2)$ at every point ζ_0 of $\mathfrak{B}$.

It should be noted that, as an immediate consequence of assumption (H. 1), there exists a first differential $J(\zeta_0; \xi)$ of I at every point ζ_0 of $\mathfrak{B}$.

We shall say that I is a minimum on $\mathfrak{B}$ at a point ζ_0 in case $I(\zeta) \geqq I(\zeta_0)$, for every ζ in $\mathfrak{B}$; we say that I is a minimum on $\mathfrak{B}_\gamma$ at a point ζ_0 in case $I(\zeta) \geqq I(\zeta_0)$, for every ζ in $\mathfrak{B}_\gamma$, which is the class of all ζ such that $\zeta - \zeta_0$ is in $\mathfrak{N}$ and ζ is in $\mathfrak{B}$.

We now state the result which is the starting point of

the theory.

LEMMA 5:1. *If I is a minimum at a point ζ_0 on $\mathfrak{Z}$ or $\mathfrak{Z}_\nu$, then*

$$J(\zeta_0;\ \delta) = 0 \quad \text{and} \quad K(\zeta_0;\ \delta,\ \delta) \geqq 0,$$

for every δ in $\mathfrak{D}$ or $\mathfrak{N}$, respectively.

It is clear that, for every δ in $\mathfrak{D}$ or $\mathfrak{N}$, $\zeta_0 + a\delta$ is in $\mathfrak{Z}$ or $\mathfrak{Z}_\nu$, respectively, for sufficiently small values of a. Whence the lemma follows immediately from Lemma 4:1.

6. *A representation of J and K.* Before deducing conditions that are necessary for I to be a minimum, we find it convenient first to prove

LEMMA 6:1. *For every ζ_0 in $\mathfrak{Z}$, there exists a regular function $j(\zeta_0; x)$ of limited variation, vanishing at 0 and 1, and a symmetric, bounded, and Riemann integrable function $k(\zeta_0; x, y)$, vanishing on the boundary of the square, such that*

$$J(\zeta_0;\ \delta) = \int_0^1 \delta(t)\,dj(\zeta_0; t) \tag{6:1}$$

$$K(\zeta_0;\ \nu,\ \nu) = \int_0^1\int_0^1 \nu'(x)k(\zeta_0; x, y)\,\nu'(y)\,dx\,dy, \tag{6:2}$$

for every δ in $\mathfrak{D}$ and ν in $\mathfrak{N}$; in addition to these properties k is of limited variation in the sense of Fréchet.

By the theorem of Riesz[1] there is a function j_0 of limited variation which is regular and vanishes at 0, and for which representation (6:1) is valid; by the defining property of $\mathfrak{D}$ it is evident that j_0 can be replaced by $j(\zeta_0; x) \equiv j_0(\zeta_0; x) - j_0(\zeta_0; 1)x$. The latter half of the theorem follows by applying Theorem 1 to K.

---

[1] See Banach, *op. cit.*, pp. 59 ff.

7. The necessary conditions. To establish certain of these conditions it is found to be convenient to make a second assumption:

(H. 2) The second differential K is of type $\mathfrak{N}$ at every point ζ_0 of $\mathfrak{Z}$.

Combining the results of the previous sections, we proceed to the demonstration of

LEMMA 7:1. A necessary and sufficient condition that

$$J(\zeta_0;\ \delta) = \int_0^1 \delta(x)\,dj(\zeta_0;\ x)$$

vanish at a point ζ_0, for every δ in $\mathfrak{D}$ or $\mathfrak{N}$, is that j is identically zero.

It suffices to show the necessity of the condition. Since J vanishes for every δ or ν, it vanishes for every ν in $\mathfrak{N}$. Then integrating by parts, we see that

$$J(\zeta_0;\ \nu) = -\int_0^1 \nu'(x)j(\zeta_0;\ x)dx = 0.$$

Hence it is clear that

$$\int_0^1 \xi(x)j(\zeta_0;\ x)dx = 0,$$

for every ξ in $\mathfrak{S}$. Consequently j is orthogonal to every Fourier function, which implies that j is identically zero. Making use of this lemma, we are in a position to state

CONDITION I. If I is a minimum at a point ζ_0 on $\mathfrak{Z}$ or $\mathfrak{Z}_\nu$, then the function j occurring in the first differential J must be identically zero on the interval from 0 to 1.

This follows at once from Lemmae 7:1 and 5:1.

Then by means of Theorem 3 and equation (6:2) the remain-

ing necessary conditions can be stated.

CONDITION II. *If I is a minimum at a point on $\mathfrak{Z}$ or $\mathfrak{Z}_\gamma$, then the lower bound B of $K(\zeta_0;\ \nu,\ \nu)$ in the class of all ν such that*

$$\int_0^1 \left[\int_0^u \nu(t)dt\right]^2 du = 1 \tag{7:1}$$

is finite; moreover B is also the lower bound of $K(\zeta_0;\ \delta,\ \delta)$ in the class of points δ which satisfy relation (7:1) with ν replaced by δ.

The first part is an immediate consequence of Lemma 6:1. The second part can be established by the continuity of K: it is clear that the lower bound B_0 of $K(\zeta_0;\ \delta,\ \delta)$ in the set of all δ such that

$$\int_0^1 \left[\int_0^u \delta(t)dt\right]^2 du = 1$$

is not greater than B. Suppose that it is less than B. Then there is a δ_0 in $\mathfrak{D}$ and a number a such that $B > a \geqq K(\zeta_0;\ \delta_0,\ \delta_0)$. Hence $K(\zeta_0;\ \nu,\ \nu) > a$; then by an argument similar to the one employed in the proof of Theorem 2 a contradiction results.

CONDITION III. *Let ζ_0 be a point at which I is a minimum on $\mathfrak{Z}$ or $\mathfrak{Z}_\gamma$ and at which $K(\zeta_0;\ \nu,\ \nu)$ is not zero for every ν in $\mathfrak{N}$; then the integral equation*

$$\int_0^1 k(\zeta_0;\ x,\ y)\,\xi(y)dy = \lambda\,\xi(x), \tag{7:2}$$

whose kernel is the continuous symmetric function appearing in representation (6:2) of K, can have no negative characteristic values; furthermore the equation always has at least one characteristic value.

It is of interest to note the following relation between the characteristic values and functions of equation (7:2) and the number B of Condition II: if λ_i is the i-th characteristic value and ξ_i a corresponding solution (normed so that the integral of its square is unity), then

$$\lambda_1 \geqq B \cdot \int_0^1 \eta_1^2(t)dt, \tag{7:3}$$

where η_1 is the solution of $\eta_1(0) = \eta_1(1) = 0$, $d^2\eta_1/dx^2 = \xi_1$.

COROLLARY 1. Let I be a minimum on $\mathfrak{Z}$ or $\mathfrak{Z}_\nu$; then

$$k_n(\zeta_0; x, x) \geqq 0 \qquad (0 \leqq x \leqq 1; n = 1, 2, \ldots),$$

where k_n is the n-th iterated kernel[1] of equation (7:2).

By the theorem of Mercer the kernel k_n can be expanded into an absolutely-uniformly convergent series

$$\sum_i \xi_i(x) \lambda_i^n \xi_i(y),$$

whenever there are no negative characteristic values λ_i;[2] this establishes the desired result.

A point ζ_0 will be said to satisfy Condition I, II, or III in case the conclusion of I, II, or III, respectively, holds at this point. It will be said to satisfy II' or III' in case the lower bound B at ζ_0 is different from zero, or every characteristic value is positive, respectively. Finally it satisfies III_N' in the event that there is an $N_d(\zeta_0)$ such that at every point ζ of this neighborhood Condition III' holds.

---

[1] The iterated kernels of an integral equation are defined as follows: $k_1(\zeta_0; x, y) \equiv k(\zeta_0; x, y)$ and

$$k_n(\zeta_0; x, y) \equiv \int_0^1 k_{n-1}(\zeta_0; x, t)k_1(\zeta_0; t, y)dt.$$

[2] See Goursat, op. cit., pp. 455 ff.

8. The sufficient conditions. These conditions will be stated in terms of the conditions of the preceding section.

LEMMA 8:1. *Let ζ_0 be a point satisfying Conditions I and III_N'. Then there is an $N_d(\zeta_0)$ for which*

$$I(\zeta_0) \leqq I(\zeta), \tag{8:1}$$

whenever ζ is in this neighborhood and in $\mathcal{B}$. Moreover when $\zeta \neq \zeta_0$ is in $\mathcal{B}_\gamma$, as well as in $N_d(\zeta_0)$, relation (8:1) holds with the equality sign excluded.

It is evident that Condition I and Lemma 7:1 imply the vanishing of the first differential at ζ_0. Moreover III_N' implies the existence of an $N_d(\zeta_0)$ such that if ζ is in this neighborhood, the second differential is non-negative at this point. Hence by means of Taylor's theorem we see that there is a number a such that $0 < a < 1$ and

$$I(\zeta) - I(\zeta_0) = K[\zeta_0 + a(\zeta - \zeta_0); \zeta - \zeta_0, \zeta - \zeta_0]/2;$$

whence the first part of the lemma is proved.

To establish the second part it suffices to note that III_N' implies that the second differential is positive at any point in $N_d(\zeta_0)$ and in $\mathcal{B}_\gamma$.

To get an analogue of the condition ordinarily stated in the calculus of variations for a weak relative minimum we make the following assumptions:

(H. 3) The lower bound $B(\zeta)$ described in Condition II is lower semi-continuous at ζ_0 on the set $\mathcal{B}$, i.e., for every $e > 0$, there is a $d > 0$ such that $B(\zeta) > B(\zeta_0) - e$ when ζ is in $N_d(\zeta_0)$ and $\mathcal{B}$.

(H. 4) The lower bound $B(\zeta)$ is lower semi-continuous at ζ_0 on $\mathcal{B}_\gamma$.

In terms of these assumptions we state

THEOREM 4. Let ζ_0 be a point at which Conditions I, II', and III' are satisfied. Further let ζ_0 satisfy (H. 3) or (H. 4). Then there is an $N_d(\zeta_0)$ with the property that

$$I(\zeta_0) < I(\zeta),$$

whenever $\zeta \neq \zeta_0$ is in this neighborhood and in $\mathfrak{Z}$ or $\mathfrak{Z}_\nu$, respectively.

Condition I implies that J vanishes at ζ_0 identically in its differential argument. Condition II' in conjunction with III' implies that $B(\zeta_0) > 0$; hence making use of (H. 3) or (H. 4), we find that $B(\zeta)$ remains positive for ζ sufficiently near to ζ_0 and in $\mathfrak{Z}$ or $\mathfrak{Z}_\nu$, respectively. Then, as before, the result follows by making use of Taylor's theorem for functionals.

9. Excursus. In this section a preliminary examination of the necessary conditions and the sufficient conditions in the light of the calculus of variations is made. However, the next section treats in considerable detail this problem. Here we are principally concerned with the Jacobi condition and its generalization to the problem at hand.

To obtain a true generalization of the Jacobi condition we consider the problem of minimizing the second differential K in the class of all ν such that

$$\int_0^1 \left[\int_0^u \nu(t)dt\right]^2 du = 1; \tag{9:1}$$

by a simple computation it can be verified that the class of all ν satisfying (9:1) is the same as the class of all ν for which

$$\int_0^1 \int_0^1 \nu'(s)q(s, t)\,\nu'(t)\,ds\,dt = 1, \tag{9:2}$$

where q is defined in the following manner:

$$q_0(s, t) \equiv st^2/2 - t^3/6 \qquad (0 \leqq t \leqq s)$$
$$\equiv s^2t/2 - s^3/6 \qquad (s \leqq t \leqq 1)$$

$$q(s, t) = q_0(s, t) + q_0(1, 1)st - q_0(s, 1)t - q_0(1, t)s.$$

The characteristic value problem associated with the problem of minimizing $K(\zeta_0; \nu, \nu)$ in the class of all ν in $\mathfrak{N}$ subject to the restriction (9:2) is evidently

$$\int_0^1 [k(\zeta_0; x, y) - \lambda q(x, y)]\nu'(y)dy = 0;$$

however, as in Section 3, it is seen that this is equivalent to the problem of solving

$$\int_0^1 [k(\zeta_0; x, y) - \lambda q(x, y)]\xi(y)dy = 0. \tag{9:3}$$

Apparently to get a generalization of the Jacobi condition from equation (9:3) it is necessary to establish two facts: first, the existence of characteristic values; and second, that these values have minimizing properties. It is our ultimate aim to show that the problem we treat does not necessarily have either of these properties. It is for this reason that Condition II was not stated in terms of characteristic values.

We shall first show that q is a positive definite kernel and hence is closed.[1] To do this it is sufficient to note that

$$\int_0^1\int_0^1 \xi(x)q(x, y)\xi(y)\,dx\,dy = \int_0^1 \eta^2(x)dx,$$

---

[1] A continuous symmetric kernel q is positive definite in case
$$\int_0^1\int_0^1 \xi(x)q(x, y)\xi(y)\,dx\,dy > 0$$
whenever ξ in $\mathfrak{C}$ is not the zero function. It is closed in case $\int_0^1 q(x, y)\xi(y)dy = 0$ implies that ξ is the zero function.

where

$$\eta(x) = \int_0^x \int_0^y \xi(z)\, dz\, dy - x\int_0^1 \int_0^y \xi(z)\, dz\, dy. \tag{9:4}$$

Further if we indicate by $\{\xi_i\}$ a normed and orthogonal basis for the equation

$$\int_0^1 q(x, y)\, \xi(y)dy = \lambda\, \xi(x), \tag{9:5}$$

we see that $\xi_1(0) = 0 = \xi_1(1)$.

Next we show that, if ξ is any element in $\mathfrak{E}$ and if η is defined by (9:4), then

$$\begin{aligned}\int_0^1 q(x, y)\, \xi(y)dy = x\int_0^x \eta(y)dy - x\int_0^1 \eta(y)dy \\ -\int_0^x y\, \eta(y)\, dy + x\int_0^1 y\, \eta(y)dy.\end{aligned} \tag{9:6}$$

It is clear that

$$\begin{aligned}\int_0^1 q_0(x, y)\, \eta''(y)dy &= \int_0^x (xy^2/2 - y^3/6)\, \eta''(y)dy \\ &\quad + \int_x^1 (x^2y/2 - x^3/6)\, \eta''(y)dy \\ &= x[(x^2/2)\, \eta'(x) - \int_0^x y\, \eta'(y)dy] - (x^3/6)\eta'(x) \\ &\quad + (x^2/2)[\eta'(1) - x\eta'(x) - \int_x^1 \eta'(y)dy] \\ &\quad + \int_0^x (y^2/2)\, \eta'(y)dy - (x^3/6)\int_x^1 \eta'(y)dy \\ &= x\int_0^x \eta(y)dy - \int_0^x y\eta(y)dy + (x^2/2 - x^3/6)\, \eta'(1);\end{aligned}$$

that

$$\int_0^1 q_0(1, 1)xy\, \eta''(y)dy = (x/3)\int_0^1 y\, \eta''(y)dy = (x/3)\eta'(1);$$

$$\int_0^1 q_0(x, 1)y\,\eta''(y)dy = (x^2/2 - x^3/6)\,\eta'(1);$$

and that

$$x\int_0^1 q_0(1, y)\,\eta''(y)dy = x\int_0^1 (y^2/2 - y^3/6)\,\eta''(y)dy$$
$$= x\left[\tfrac{1}{3}\eta'(1) + \int_0^1 \eta(y)dy - \int_0^1 y\,\eta(y)dy\right].$$

An example will now be given of a continuous, symmetric function k of limited variation in the sense of Fréchet which vanishes on the boundary of the square and which has the property that equation (9:3) has no characteristic values.

We define k as follows:

$$k_0(s, t) \equiv t^5/120 - st^4/24 + s^2t^3/12 \qquad (0 \leqq t \leqq s)$$
$$\equiv s^5/120 - s^4t/24 + s^3t^2/12 \qquad (s \leqq t \leqq 1)$$
$$k(\zeta_0; s, t) \equiv k_0(s, t) + k_0(1, 1)st - k_0(s, 1)t - k_0(1, t)s.$$

By a direct computation making use of the definitions of k, k_0, of relation (1:4), and of integration by parts, one can verify that

$$\int_0^1 k(\zeta_0; s, t)\,\xi(t)dt = \int_0^s (t^3/6 - st^2/2 + s^2t/2)\,\eta(t)dt$$
$$- s\int_0^1 (t^3/6 - t^2/2 + t/2)\,\eta(t)dt$$
$$+ \frac{s^3}{6}\int_s^1 \eta(t)dt.$$

Suppose that, for this choice of k, equation (9:3) has a characteristic value λ_0 and a characteristic solution ξ_0, which is not identically zero; then η_0, defined by relation (9:4), is not identically zero. Moreover

$$
\begin{aligned}
&\int_0^1 [k(\zeta_0; s, t) - \lambda_0 q(s, t)]\, \xi_0(t)dt \\
&\quad = \int_0^1 [k(\zeta_0; s, t) - \lambda_0 q(s, t)]\, \eta_0''(t)dt \\
&\quad = \int_0^s (t^3/6 - st^2/2 + s^2 t/2)\, \eta_0(t)dt \\
&\quad\quad - s\int_0^1 (t^3/6 - t^2/2 + t/2)\, \eta_0(t)dt \\
&\quad\quad + \frac{s^3}{6}\int_s^1 \eta_0(t)dt - \lambda_0 s\int_0^s \eta_0(t)dt \\
&\quad\quad + \lambda_0 \int_0^s t\, \eta_0(t)dt + \lambda_0 s\int_0^1 \eta_0(t)dt \\
&\quad\quad - \lambda_0 s\int_0^1 t\, \eta_0(t)dt = 0
\end{aligned}
\tag{9:5}
$$

Differentiating both members three times, we get the following equation

$$\int_s^1 \eta_0(t)dt - \lambda_0 \eta_0'(s) = 0,$$

which implies that $\lambda_0 = 0$ or $\eta_0'(1) = 0$. However if λ_0 were zero, then η_0 would be identically zero, which is a contradiction; therefore $\eta_0'(1) = 0$. Differentiating still again, we are led to the following differential equation:

$$\lambda_0 \eta_0''(s) + \eta_0(s) = 0$$
$$\eta_0(0) = 0 = \eta_0(1) = \eta_0'(1).$$

But the only solution this equation has is the zero function. Hence we have given an example for which equation (9:3) has no characteristic values. Next we show that even when characteristic values exist, they do not necessarily have minimizing properties. Let $k(\zeta_0; x, y)$ be defined to be $\sum_n \xi_n(x)\, \xi_n(y)/(n\pi)^8$, where $\xi_n(x) = \sqrt{2} \sin n\pi x$. Then k is symmetric, vanishes on the

boundary of the square, is continuous, and has a continuous mixed partial derivative; in fact, it is not difficult to show that $\partial^4 k(\zeta_0; x, y)/\partial x^2 \partial y^2 = q(x, y)$. Moreover by the remark immediately preceding Corollary 1 of Section 7, it is seen that

$$1/(n\pi)^8 \geqq B/(n\pi)^4 \qquad (n = 1, 2, \ldots),$$

since the characteristic values of the integral equation having k as its kernel are $1/(n\pi)^8$. Hence $B = 0$, since in this example it is never negative. Therefore no characteristic value is the minimum value of K in the class (9:2).

CHAPTER IV

APPLICATIONS OF THE GENERAL THEORY

Introduction. In the first section of this chapter the ordinary problem of the calculus of variations with fixed end-points in the plane is shown to be a special case of the theory set forth in the previous chapter. A detailed study of the conditions obtained is made for this special case, and their relations to the conditions of the calculus of variations is shown. In the next section it is shown that there are non-calculus of variations problems which fall under this theory; in particular one such special case is examined in some detail.

10. The calculus of variations. We proceed to an interpretation of our results for the fixed end-point problem of the calculus of variations. In this domain the functional I is of the form

$$I(\zeta) = I(y_0, y') = \int_0^1 f(x, y, y')dx,$$

where $y_0 = y(0)$, $\zeta = y'$. This functional satisfies the assumptions made in the previous sections.

One sees at once that the first variation of I is expressible as

$$\int_0^1 \eta'(x)d\left\{\int_0^x [f_{y'}(s, y, y') - \int_0^s f_y(t, y, y')dt]ds\right\}.$$

Hence Condition I implies that

$$f_{y'} - \int_0^x f_y dx = a;$$

which upon differentiation leads to the Euler equation.

To transform the second variation into a double Fréchet-Stieltjes integral the following definitions will be made:

$$P_1 \equiv f_{yy}, \qquad P_2 = f_{yy'}, \qquad P_3 = f_{y'y'};$$

$$p_i(s, t) \equiv \int_0^t P_i(u)du \qquad (0 \leqq t \leqq s;\ i = 1, 2, 3)$$

$$\equiv \int_0^s P_i(u)du \qquad (s \leqq t \leqq 1);$$

$$q_1(s, t) \equiv \int_0^s du \int_0^t p_1(u, v)dv + st\int_0^1\int_0^1 p_1(u, v)\, du\, dv$$

$$- t\int_0^s du\int_0^1 p_1(u, v)dv - s\int_0^1 du\int_0^t p_1(u, v)dv \tag{10:1}$$

$$h(s, t) \equiv \int_0^s p_2(u, t)du + st\int_0^1 p_2(u, 1)du$$

$$- t\int_0^s p_2(u, 1)du - s\int_0^1 p_2(u, t)du$$

$$q_2(s, t) \equiv [h(s, t) + h(t, s)]/2$$

$$q_3(s, t) \equiv p_3(s, t) + stp_3(1, 1) - sp_3(1, t) - tp_3(s, 1)$$

$$k(s, t) \equiv q_1(s, t) - 2q_2(s, t) + q_3(s, t).$$

In terms of these definitions it can be verified that the second variation of I is expressible as

$$\int_0^1\int_0^1 \eta'(s)\, \eta'(t)d_s d_t k(s, t).$$

It is clear that the number B appearing in Condition II is equal to the least characteristic value for the boundary value problem normally associated with the second variation, provided one exists. In fact our Condition II is equivalent to the Legendre condition since if $f_{y'y'}$ is less than zero at a point, the bound $B = -\infty$.

Observing that

$$\frac{\partial^2 q_1(s,\,t)}{\partial t^2} = \frac{\partial}{\partial t}\left[\int_0^s p_1(u,\,t)\,du - s\int_0^1 p_1(u,\,t)\,du\right]$$

$$= (s - t)P_1(t) - s(1 - t)P_1(t) \qquad (0 \leqq t \leqq s)$$

$$= -s(1 - t)P_1(t) \qquad (s \leqq t \leqq 1)$$

$$\frac{\partial^2 q_2(s,\,t)}{\partial t^2} = \frac{1}{2}\,\frac{\partial}{\partial t}\left\{\frac{\partial}{\partial t}\left[\int_0^s p_2(u,\,t)\,du - s\int_0^1 p_2(u,\,t)\,du\right] + p_2(s,\,t) - sp_2(1,\,t)\right\}$$

$$= \frac{1}{2}[(s - t)P_2'(t) - s(1 - t)P_2'(t)] \qquad (0 \leqq t \leqq s)$$

$$= \frac{1}{2}[-s(1 - t)P_2'(t)] \qquad (s \leqq t \leqq 1),$$

we see that the integral equation which appears in Condition III becomes

$$-\int_0^s \eta'(t)P_3(t)\,dt + \int_0^s \eta(t)[-(s - t)P_2'(t) + (s - t)P_1(t)]\,dt$$

$$- s\int_0^1 \eta(t)[P_3'(t) - (1 - t)P_2'(t) + (1 - t)P_1(t)]\,dt = \lambda\eta''(s).$$

It is apparent that the left member of this equation has a continuous second derivative; hence, differentiating both members and remembering that $k(s,\,t)$ vanishes on the boundary of the square, we are led to the differential equation

$$(10{:}2) \qquad \frac{d}{ds}[\eta'(s)P_3(s)] + \eta(s)[P_2'(s) - P_1(s)] = -\lambda\eta^{iv}(s)$$

with the boundary conditions

$$(10{:}3) \qquad \eta(0) = \eta(1) = 0 = \eta''(0) = \eta''(1)$$

when λ is different from zero and with

$$(10{:}4) \qquad \eta(0) = 0 = \eta(1)$$

when λ is equal to zero. Whence Condition III becomes

THEOREM 5. For every admissible arc E which gives I a minimum there can exist no negative values λ which are characteristic for equation (10:2) with boundary conditions (10:3). Moreover if the second variation is not identically zero, then there always exists at least one non-zero characteristic value.

In an analogous fashion the remark preceding Corollary 1 of Condition III and the corollary itself can be interpreted for the calculus of variations.[1]

It is of interest to note the relationships between the Jacobi condition and Condition III, as stated in Theorem 5. In the first place it is clear that even if $f_{y'y'}$ is not always different from zero, there exists a continuous solution of equation (10:2); this is not necessarily true for the accessory boundary value problem since the ordinary existence theorems for differential equations are not applicable at points where $f_{y'y'} = 0$. In fact the accessory minimum problem associated with the problem of minimizing

$$\int_0^1 xy^2 dx$$

in the class of arcs joining (0, 0) to (1, 0) is $-x\eta(x) = \lambda\eta(x)$; clearly this equation has no characteristic values, and $f_{y'y'} = 0$ for every x. For this problem equation (10:2) becomes $+x\eta(x) = +\lambda\eta^{iv}(x)$, $\eta(0) = 0 = \eta(1)$ and $\eta''(0) = 0 = \eta''(1)$ since $\lambda = 0$ cannot be characteristic. By our general theory this equation has at least one characteristic value $\lambda \neq 0$. Moreover, if $\lambda = 0$ is characteristic for equation (10:2) with boundary conditions

---

[1] It seems that this remark and corollary do not appear in the literature of the calculus of variations; apparently Condition III does not either.

(10:4) it is characteristic for the accessory boundary value problem and conversely. On the other hand if every characteristic value of one is positive, so is every characteristic value of the other.

The strengthened Legendre and Jacobi conditions imply our Condition III_N', but the converse is not true; to show this it suffices to give as an example the problem of minimizing

$$I = \int_0^1 y^2(x)dx \tag{10:5}$$

in the class of arcs joining (0, 0) to (1, 0). It is evident that $f_{y'y'} = 0$ and that $k(\zeta_0; s, t)/2$ is the function $q(s, t)$ defined in Section 9. Then the equation

$$\int_0^1 k(\zeta_0; s, t)\xi(t)dt = \lambda\xi(s).$$

has only positive characteristic values. Moreover since the second differential of I is independent of y, k is also; hence this equation is valid for ζ_0 replaced by any ζ. Consequently Condition III_N' is satisfied, but the strengthened Legendre condition fails.

For the problem of minimizing

$$I = \int_0^1 y'^2(x)dx$$

we note that equation (10:2) becomes $2\eta''(s) = -\lambda\eta^{iv}(s)$, $\eta(0) = 0 = \eta(1) = \eta''(0) = \eta''(1)$. Moreover $\lambda = 0$ cannot be characteristic. Hence $\lambda_m = 2/m^2\pi^2$ (m = 1, 2, ...) are the characteristic values.

It is clear from what has been said in this section that Condition III, as stated in Theorem 6, is applicable to a wider

class of problems than the ordinary Jacobi condition which requires that the arc be non-singular; that is, $f_{y'y'} \neq 0$ at every point on the arc in question. Moreover the condition does not require that the arc have a second derivative. Furthermore the problem stated in (10:5) shows that both Condition II' and Condition III_N' can be satisfied without the strengthened Legendre condition holding. Hence the sufficiency theorems of Section 9 are more general than the standard sufficiency theorem for a weak relative minimum since they have weaker hypotheses.

11. A non-calculus of variations example. To show that the theory which has been formulated in the preceding pages is more general than the ordinary problem of the calculus of variations it suffices to give a non-calculus of variations instance of this theory. One such example is the problem of minimizing

$$\int_0^1 f(x, y, y')dg(x)$$

in the class of arcs joining two fixed points $(0, y_1)$ and $(1, y_2)$, where f has the properties normally assumed in the calculus of variations and g is of limited variation. Another example of greater interest is the problem of minimizing u(1), where u(x) is determined by the equation

$$u(x) = \int_0^x f[x, t, u(t), y(t), y'(t)]dt.$$

However, we shall content ourselves with examining in some detail the problem of minimizing

$$I(\zeta) = I(y_1, y') = \int_0^1 \int_0^1 g[s, t, y(s), y'(s), y(t), y'(t)]\, ds\, dt$$

in the class of arcs joining $(0, y_1)$ to $(1, y_2)$. For this purpose

it will be supposed that $g(s, t, u, v, w, z)$ is of class C^{iv}. Moreover, the following notations will be introduced to simplify the computations:

$$p_1(s) \equiv \int_0^1 g_u[s, t, y(s), y'(s), y(t), y'(t)]dt + \int_0^1 g_w[t, s, y(t), y'(t), y(s), y'(s)]dt$$

$$p_2(s) \equiv \int_0^1 g_v[s, t, y(s), y'(s), y(t), y'(t)]dt + \int_0^1 g_z[t, s, y(t), y'(t), y(s), y'(s)]dt$$

$$q_1(s) \equiv \int_0^1 g_{uu}[s, t, y(s), y'(s), y(t), y'(t)]dt + \int_0^1 g_{ww}[t, s, y(t), y'(t), y(s), y'(s)]dt$$

$$q_2(s) \equiv \int_0^1 g_{uv}[s, t, y(s), y'(s), y(t), y'(t)]dt + \int_0^1 g_{wz}[t, s, y(t), y'(t), y(s), y'(s)]dt$$

$$q_3(s) \equiv \int_0^1 g_{vv}[s, t, y(s), y'(s), y(t), y'(t)]dt + \int_0^1 g_{zz}[t, s, y(t), y'(t), y(s), y'(s)]dt$$

$$r_1(s, t) \equiv g_{uw}[s, t, y(s), y'(s), y(t), y'(t)] + g_{uw}[t, s, y(t), y'(t), y(s), y'(s)]$$

$$r_2(s, t) \equiv g_{uz}[s, t, y(s), y'(s), y(t), y'(t)] + g_{wv}[t, s, y(t), y'(t), y(s), y'(s)]$$

$$r_3(s, t) \equiv g_{vz}[s, t, y(s), y'(s), y(t), y'(t)] + g_{vz}[t, s, y(t), y'(t), y(s), y'(s)].$$

In terms of these definitions and Lemma 4:1 it is clear that the first differential of I is

$$\int_0^1 [p_1(s)\,\eta(s) + p_2(s)\,\eta'(s)]ds = \int_0^1 \eta'(s)dj(s),$$

where

$$j(s) \equiv \int_0^s [p_2(t) - \int_0^t p_1(u)du]dt - s\int_0^1 [p_2(t) - \int_0^t p_1(u)du]dt; \tag{11:1}$$

and that the second differential is

$$\int_0^1 [q_1(s)\,\eta^2(s) + 2q_2(s)\,\eta(s)\,\eta'(s) + q_3(s)\,\eta'^2(s)]ds + \int_0^1\int_0^1 [\eta(s)r_1(s, t)\,\eta(t) + 2\,\eta(s)r_2(s, t)\,\eta'(t) + \eta'(s)r_3(s, t)\,\eta'(t)]ds\,dt. \tag{11:2}$$

To transform this expression into a double Fréchet-Stieltjes integral we proceed in the following fashion. Replacing the functions P_1 in equations (10:1) by $q_i(s)$ $(i = 1, 2, 3)$, we define a function k_0 such that the single integral in (11:2) becomes

$$\int_0^1\int_0^1 \eta'(s)\,\eta'(t)d_sd_tk_0(s, t).$$

The terms in the double integral appearing in expression (11:2) become, after integrating by parts

$$\int_0^1\int_0^1 \eta'(s)\,\eta'(t)d_sd_tk_i(s, t) \qquad (i = 1, 2, 3),$$

where

$$k_1(s, t) \equiv \int_0^s dx\int_0^t dy\int_0^x dz\int_0^y r_1(z, w)dw$$

$$k_2(s, t) \equiv -\int_0^s dx\int_0^t dy\int_0^x r_2(z, y)dz - \int_0^s dy\int_0^t dx\int_0^x r_2(z, y)dz$$

$$k_3(s, t) \equiv \int_0^s dx\int_0^t r_3(x, y)dy.$$

Hence the functions k_i are symmetric. Finally we define k by the relation

$$k(\zeta_0; s, t) \equiv \sum_{i=0}^{3} [k_i(s, t) + k_i(1, 1)st - k_i(s, 1)t - k_i(1, t)s].$$

Thus expression (11:2) becomes

$$\int_0^1 \int_0^1 \eta'(s)\, \eta'(t) d_s d_t k(\zeta_0; s, t),$$

and the equation

$$\int_0^1 k(\zeta_0; s, t)\, \eta''(t) dt = \lambda \eta''(s)$$

becomes

$$- \int_0^s \eta'(t) q_3(t) dt + \int_0^s \eta(t) [-(s - t) q_2'(t) + (s - t) q_1(t)] dt$$
$$- s \int_0^1 \eta(t) [q_3'(t) - (1 - t) q_2'(t) + (1 - t) q_1(t)] dt$$
$$+ \sum_{i=1}^{3} \int_0^1 \frac{\partial^2}{\partial t^2} [k_i(s, t) - s k_i(1, t)]\, \eta(t) dt = \lambda \eta''(s).$$

Then, as in Section 10 we differentiate both sides twice with respect to s and get

$$\frac{d}{ds} [\eta'(s) q_3(s)] + \eta(s) [q_2'(s) - q_1(s)]$$
$$- \sum_{i=1}^{3} \int_0^1 \frac{\partial^4 k_i(s, t)}{\partial s^2\, \partial t^2} \eta(t) dt = - \lambda \eta^{iv}(s), \tag{11:3}$$

where $\eta(0) = 0 = \eta(1)$ and if $\lambda \neq 0$, $\eta''(0) = 0 = \eta''(1)$. We recall from the general theory that if the second differential is not identically zero, there is always a non-zero characteristic value λ.

Two of the necessary conditions we deduced state that the function j appearing in (11:1) is identically zero, and that equa-

tion (11:3) can have no negative characteristic values. The other conditions and results of Sections 7 and 8 can similarly be interpreted for this situation.